DOCUMENT RESUME

ED 211 580 TM 810 970

AUTHOR Davison, Mark L.
TITLE Multidimensional Scaling vs. Factor Analysis of Tests and Items.
PUB DATE 26 Aug 81
NOTE 23p.: Paper presented at the Annual Meeting of the American Psychological Association (Los Angeles, CA, August 26, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Ability; *Comparative Analysis; Correlation; *Factor Analysis; Factor Structure; *Multidimensional Scaling; *Psychological Testing; Test Interpretation; Test Items; Vocational Interests
IDENTIFIERS General Aptitude Test Battery; Vocational Preference Inventory

ABSTRACT

Academic psychology has long been composed of two disciplines, one experimental and one correlational. These two disciplines each developed their own method of studying structure in data: multidimensional scaling (MDS) and factor analysis. Both methods use similar kinds of input data, proximity measures on object pairs. Both represent the object structure in terms of spatial coordinates. When MDS and factor analysis are applied to the same test intercorrelation matrix, how do the results compare? In an analysis of ability data and an analysis of vocational interest data, two-dimensional, nonmetric MDS solutions were compared to three-factor, principal components solutions. In both analyses, the components solution contained a general factor with no counterpart among the scaling dimensions. Loadings along the remaining two components closely resembled scale values along the two dimensions. Results suggest that if one compares a K-dimensional MDS solution to a (K+1) components analysis, the components analysis will often contain a general factor with no counterpart among the scaling dimensions; after applying an appropriate rotation and multiplicative constant to the MDS scale values, some or all of the remaining components will correspond to a dimension in the scaling solution. (Author)

Multidimensional Scaling vs. Factor Analysis of Tests and Items

Mark L. Davison

University of Minnesota

Address presented to Division 5 of the American Psychological Association, Los Angeles, California, August 26, 1981.

Abstract

Academic psychology has long been composed of two disciplines, one experimental and one correlational. These two disciplines each developed their own method of studying structure in data: multidimensional scaling (MDS) and factor analysis. Both methods use similar kinds of input data, proximity measures on object pairs. Both represent the object structure in terms of spatial coordinates. When MDS and factor analysis are applied to the same test intercorrelation matrix, how do the results compare? In an analysis of ability data and an analysis of vocational interest data, two dimensional, nonmetric MDS solutions were compared to three-factor, principal components solutions. In both analyses, the components solution contained a general factor with no counterpart among the scaling dimensions. Loadings along the remaining two components closely resembled scale values along the two dimensions. Results suggest that if one compares a K-dimensional MDS solution to a (K + 1) components analysis, the components analysis will often contain a general factor with no counterpart among the scaling dimensions; after applying an appropriate rotation and multiplicative constant to the MDS scale values, some or all of the remaining components will correspond to a dimension in the scaling solution.

Multidimensional Scaling vs. Factor Analysis of Tests and Items

Mark L. Davison

University of Minnesota

Twenty-four years ago, Lee Cronbach delivered his Presidential address to the American Psychological Association, an address entitled "The two disciplines of scientific psychology." In that address, Cronbach pointed to a schism in academic psychology between the experimental and correlational traditions. Cronbach described the experimentalists as psychologists who employ the experimental method to bring variables under tight control. This control permits "rigorous tests of hypotheses and confident statements about causation." (Cronbach, 1967, p.23) In contrast to the experimentalists, the correlational psychologists "study what man has not learned to control or can never hope to control".

Not only did the two disciplines develop their own measurement methods, each developed its own statistical technique for studying structure. Correlational psychology developed factor analysis. Multidimensional scaling is largely a creature of experimental psychology.

The main purpose of my talk today is to compare factor analysis and multidimensional scaling as statistical techniques for studying the structure of psychological tests and test items. Specifically, I want to illustrate how a factor analysis of test (or item) intercorrelations often compares to a multidimensional scaling of the same intercorrelations. Before making this comparison, however, permit me a small diversion into a discussion of parallels between the two techniques. These parallels have made factor analysis and multidimensional scaling rival alternatives for the study of structure in psychology.

Parallels between Factor Analysis and Multidimensional Scaling

There are at least three kinds of parallels between factor analysis and multidimensional scaling; parallels in their historical developments, parallels in the way they represent structure, and parallels in the purposes which they have served. Let's consider each of these parallels in turn beginning with the historical.

Neither factor analysis nor multidimensional scaling began as statistical techniques per se. Both methods grew out of attempts to estimate the parameters in some psychological theory. Only later did these parameter estimation techniques manage to separate themselves from the psychological theories of their origins to become statistical theories in their own right. Before factor analysis, there were factor theories of human abilities, such as that of Spearman (1904), Vernon (1950), and Thurstone (1938). Factor analysis began as a method for estimating the parameters in these factor theories. After awhile, it became a statistical technique applicable not just to the study of human abilities, but to phenomena in all the behavioral, social, and natural sciences.

Similarly, before there was multidimensional scaling, Richardson (1938) proposed a distance model for psychophysical judgments of similarity between pairs of stimuli. Young and Householder (1938) developed a method to estimate the parameters in Richardson's psychophysical model. Other people expanded on their parameter estimation method until it developed into what we now call multidimensional scaling.

The parallels between the two methods go beyond the historical ones. Both techniques use the same kind of input data, measures of proximity on pairs of objects. The correlation coefficient is the overwhelmingly favored proximity measure for the factor analysts. Although multidimensional scaling

advocates do not so consistently favor one proximity measure, nevertheless, they like the factor analysts must use some measure of proximity or association between object pairs as the input to their analysis.

Not only do the methods use comparable input data, they yield parallel representations of structure. That is, both techniques yield a representation of structure in terms of spatial coordinates. Factor analysts call their coordinates 'factor loadings' and multidimensional scaling advocates call their coordinates 'scale values.'

Given the above parallels, it is not surprising to find that the two methods have been used to study many of the same issues in psychology. Social psychologists have used both factor analysis and multidimensional scaling to study the social dimensions underlying person perception (Jone & Young, 1972; Rosenberg & Jones, 1972; Taguiri, 1958). Industrial/organizational psychologists have used the two techniques to study dimensions of job performance (Borman, Hough, & Dunnette, 1976; Smith & Siegel, 1967). Educational psychologists have used the two to study the structure of human abilities (Schlessinger & Guttman, 1969).

When factor analysis and multidimensional scaling are used to study the same issue, how do the results compare? Particulary, when the two methods are used to study the intercorrelations of tests, how does the factor structure compare to the multidimensional scaling?

Prior Comparisons of Multidimensional Scaling and Factor Analysis

Other people have compared multidimensional scaling and factor analysis (MacCallum, 1974; Schlessinger & Guttman, 1969; Shepard, 1972). One conclusion appears several times in these comparisons; multidimensional scaling solutions tend to be simpler than factor solutions. That is, the number of dimensions in scaling solutions tends to be less than the number of factors in factor solutions.

Shepard (1972) has concluded that multidimensional scaling solutions tend to be simpler in the manner described. In his comparison of the two techniques, he emphasizes that in nonmetric multidimensional scaling, a small number of dimensions often suffices to represent the structure of the data. Further, he notes that ten or more factors are extracted in some factor analytic studies. "Such results cannot, of course, be cast into a readily visualizable picture. This matter of dimensionality and hence visualizability tends, in practice, to distinguish these relatively new methods of multidimensional scaling from the related methods that have long been used in the social sciences under such names as 'factor analysis' and 'principal components analysis'." (Shepard, 1972, pp. 2,3) Schlessinger and Guttman (1969) draw much the same conclusion.

If multidimensional scaling solutions really are simpler than factor solutions, then how is this simplicity achieved? Do the scaling dimensions somehow represent all of the important structure in a set of tests? Or do the scaling dimensions oversimplify the structure of the tests (or test items) by omitting important features of the structure? If the latter is true, then the more complex factor structure may better represent the relationships between the tests than does the simpler, multidimensional scaling solutions.

The next two sections of this paper present factor and multidimensional scaling analyses of the same data so that we can compare the two. In comparing them, we should keep in mind the broad question: what is the relationship between a multidimensional scaling and factoring of the same test intercorrelations? We should also keep in mind a second question raised by Shepard's (1972) work:

If the multidimensional scaling is simpler, is that simplicity achieved, in part, by omitting important features of the test structure which are contained in the factor loadings?

Human Ability Tests

Let's begin the empirical comparison of factor analysis and multidimensional scaling by examining some data from the field in which factor analysis has its roots, the field of human abilities. For the ability data to be shown below, there is an interesting correspondence between the three-dimensional, principal components solution and the two-dimensional, nonmetric multidimensional scaling solution.

Table 1 shows the intercorrelations of twelve subtests from the General Aptitude Test Battery (United states Government Printing Office, 1970) published by the U.S. Department of Labor. The subjects were 168 clients in the Vocational Assistance Program of the Minnesota Department of Vocational Rehabilitation.

Table 2 shows the first three unrotated components from a principal components analysis of these data. It also shows the scale values from a two-dimensional, nonmetric analysis of these same correlations (Kruskal, Young, & Seery, 1973).

The first principal component is a genreal factor along which every subtest has a relatively high, positive loading. It has no counterpart among the two scaling dimensions. That is, there is no scaling dimension on which all tests have high, positive scale values. Both scaling dimensions are bipolar; some tests have positive scale values and some tests have negative ones.

Unlike the first principal component, the second component in these data does have a counterpart among the scaling dimensions. If you compare each factor loading along Factor II to the corresponding scale value along Dimension I, you

Table 1

Pearson Product-Moment Correlation Coefficients*

for the General Aptitude Test Battery

Subtest	1	2	3	4	5	6	7	8	9	10	11	12
1 NAMES		697	360	637	586	552	496	561	338	349	390	354
2 ARITH			366	580	471	760	411	501	297	247	319	325
3 3-D				528	554	468	580	249	276	279	358	234
4 VOCAB					425	616	444	465	211	209	267	283
5 TOOLS						369	531	444	292	336	361	267
6 MATH							400	407	300	234	208	311
7 SHAPES								387	323	401	444	428
8 MARKING									494	540	439	422
9 PLACE										773	468	453
10 TURN											476	482
11 ASMBL												676
12 DISASMBL												

*decimal points deleted. I am indebted to Stephen Prestwood for bringing these correlations to my attention.

Table 2

Factor Loadings and Nonmetric Scale Values

for the General Aptitude Test Battery

Subtest	Factor I	Dimension I	Factor II	Dimension II	Factor III
NAMES	.785	-.191	-.255	-.129	-.191
ARITHMETIC	.742	-.349	-.381	-.277	-.331
3-D	.635	-.385	-.232	.403	.564
VOCABULARY	.700	-.476	-.436	-.101	-.059
TOOLS	.691	-.274	-.171	.253	.316
MATH	.694	-.435	-.422	-.284	-.257
SHAPES	.712	-.079	-.032	.329	.440
MARKING	.723	.147	.150	-.264	-.327
PLACE	.620	.566	.546	-.281	-.172
TURN	.633	.527	.585	-.161	-.085
ASSEMBLE	.646	.415	.437	.280	.184
DISASSEMBLE	.624	.535	.433	.233	.006

will see that every one of those loadings is of the same sign and approximately the same magnitude as the corresponding scale value.

Similarly, the third principal component has its counterpart among the scaling dimensions. If you compare each loading along Factor III to the corresponding scale value along Dimension II, you will see that all of the loadings are of the same sign and of approximately the same magnitude as the corresponding scale value.

One can compute a coefficient of congruence for the two dimensions and the corresponding two factors. In the present case, that congruence coefficient is simply the correlation between the 24 scale values along Dimensions I and II and the corresponding 24 loadings along Factors II and III. The congruence coefficient equals .96 and indicates that there is a high degree of correspondence between the dimensions and the second and third principal components.

Figure 1 graphically displays the relationship shown in Table 2 between Dimensions I and II of the scaling solution and Factors II and III of the components analysis. In this figure, squares represent subtest scale values. For each square, the coordinate along the horizontal axis represents the Dimension I scale value of the corresponding subtest. The coordinate along the vertical axis is the Dimension II scale value. Circles in Figure 2 represent subtest factor loadings. The coordinate along the horizontal axis is the Factor II loading for the corresponding subtest. The coordinate along the vertical axis is the Factor III loading. For each test, an arrow connects the square representing its scale values to the corresponding circle representing its factor loadings. This graph shows that the Factor II and III loadings (the circle) place each subtest in approximately the same region of the space as do its Dimension I and II scale values (the square).

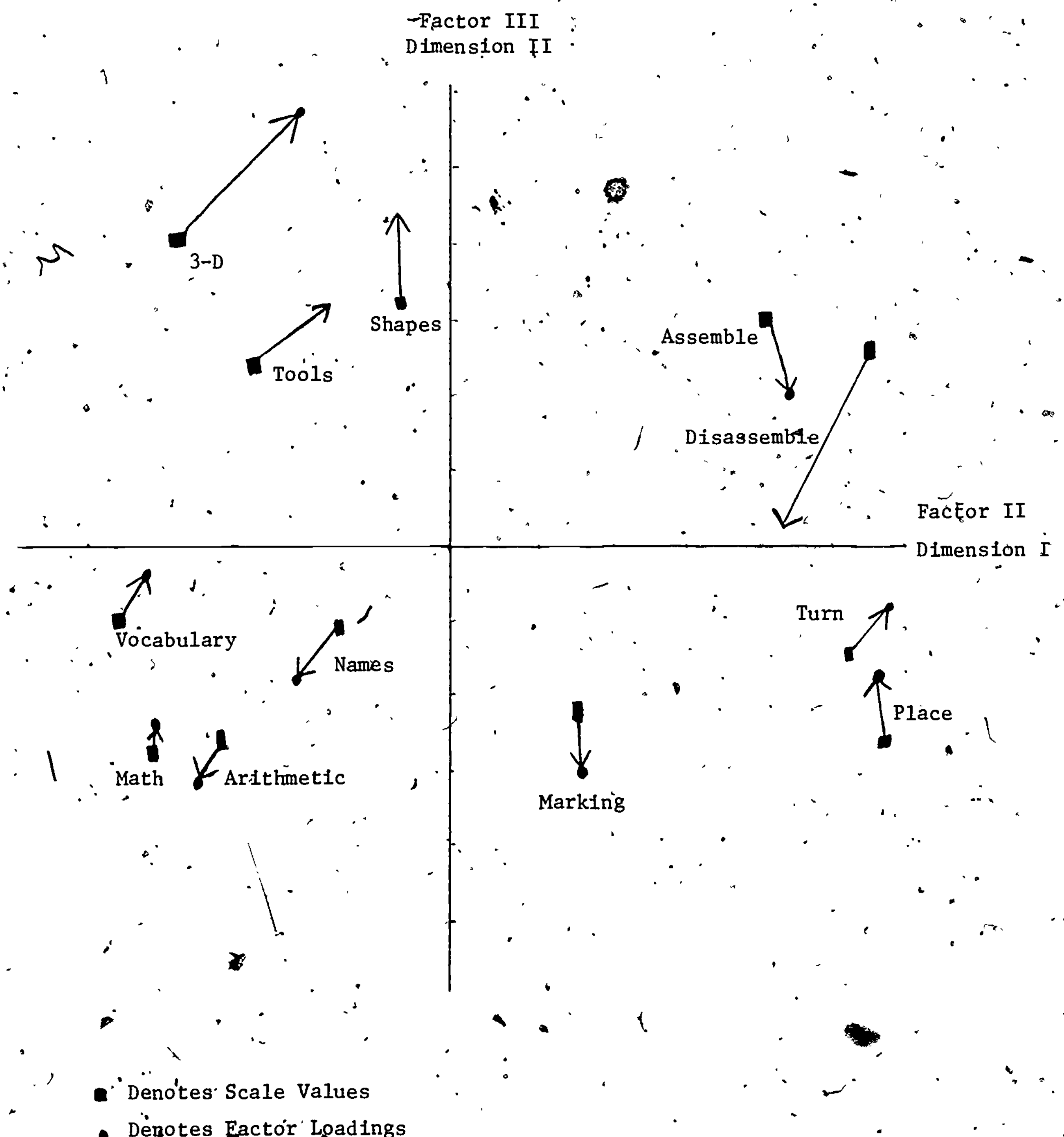

Figure 1: Scale Values and Factor Loadings for the General Aptitude Test Battery.

Table 2 and Figure 1 illustrate what may be a common, but not universal, relationship between a multidimensional scaling and factor analysis of the same data. If one compares a (K + 1)-dimensional factor analysis to a K-dimensional multidimensional scaling solution, then with a proper rotation of the two solutions, one can find a general factor which has no counterpart in the scaling solution; beyond the general factor, however, each major factor will have a counterpart among the scaling dimensions.

Occupational Interest Tests

Outside of the area from which factor analysis emerged--the study of human abilities--the general factor has not always played such a theoretically prominent role. In theories of personality, attitudes, and interests, where the supporting data are heavily based on self-report questionnaires, the general factor has sometimes been dismissed as a theoretically uninteresting, response level factor. In fact one can find factor studies in which the general factor isn't even reported. Consider this quotation from a study of self-reported occupational interests by Hanson, Prediger, and Schussel (1977):

> Not shown... is a general factor common to interest inventories using response categories such as "like," "indifferent," and "dislike". When such categories are used, the frequency with which a particular response is chosen tends to vary from person to person, regardless of item content. That is, some persons tend to choose "indifferent" more often, etc. Hence, there is a general response-related factor affecting the scores on each scale. The chief identifying feature of this factor, to be called the "response level factor," is that all interest scales have relatively high loadings on it. (Hanson et al., 1977, p. 20, italics added)

As the first line of this quotation indicates, these author accorded so little import to the general factor that they decided to omit it from their tables of factor loadings. As we shall see below, it is precisely this general factor

Table 3

Intercorrelations of the Vocational Preference Inventory Scales for a Sample of 1234 men

	R	I	A	S	E	C
REALISTIC	1.00	.46	.16	.21	.30	.36
INVESTIGATIVE	.46	1.00	.34	.30	.16	.16
ARTISTIC	.16	.34	1.00	.42	.35	.11
SOCIAL	.21	.30	.42	1.00	.54	.38
ENTERPRISING	.30	.16	.35	.54	1.00	.68
CONVENTIONAL	.36	.16	.11	.38	.68	1.00

which is missing in a multidimensional scaling solution based on data of the kind analyzed by Hanson et al.

Table 3 shows the kind of data Hanson et al. analyzed in the report from which the above quote was taken. These are the intercorrelations of six subtests from Holland's (1965) Vocational Preference Inventory (VPI). The

VPI contains six scales, each of which corresponds to a type of occupation in Holland's (1973) theory of careers. The six scales, which are named after the six occupational types, are called the Realistic, Investigative, Artistic, Social, Enterprising, and Conventional scales. Table 3 shows the intercorrelations of the six scales in a sample of 1234 men (Holland, Whitney, Cole, & Richards, 1969). This is one of five such correlation matrices analyzed in a study by Rounds, Davison, and Dawis (1979).

Table 4 shows the first three unrotated components from a principal components analysis of these data. It also shows the scale values from a two-dimensional, nonmetric analysis of these same correlations (Kruskal, Young, & Seery, 1973).

The first principal component is the general factor to which the quotation above refers, and it displays the distinctive feature described in that quotation--"all interest scales have relatively high loadings on it." (Hanson et al., 1977, p. 20) This general factor, which Hanson _et al._ (1977) chose not to report, has no counterpart in the multidimensional scaling dimensions. That is, there is no scaling dimension on which all tests have high, positive scale values. Both scaling dimensions are bipolar; some tests have positive scale values and some tests have negative ones.

Unlike the first principal component, the second component in these data does have a counterpart among the scaling dimensions. If you compare each loading along Factor II to the corresponding scale value along Dimension I, you will see that every one of those loadings is of the same sign and approximately the same magnitude as the corresponding scale value.

Similarly, the third principal component has its counterpart among the scaling dimensions. If you compare each loading along Factor III to the corresponding scale value along Dimension II, you will see that all but one of the loadings are of the same sign and approximately the same magnitude as

Factor Loadings and Nonmetric Scale Values

for the Six Vocational Preference Inventory

Scales

SCALE	Factor I	Dimension I	Factor II	Dimension I	Factor III
REALISTIC	.595	.369	.324	.491	.598
INVESTIGATIVE	.555	.683	.685	.061	.158
ARTISTIC	.568	.133	.335	-.696	-.607
SOCIAL	.736	-.236	-.080	-.294	-.371
ENTERPRISING	.803	-.492	-.437	.020	-.055
CONVENTIONAL	.711	-.456	-.496	.418	.307

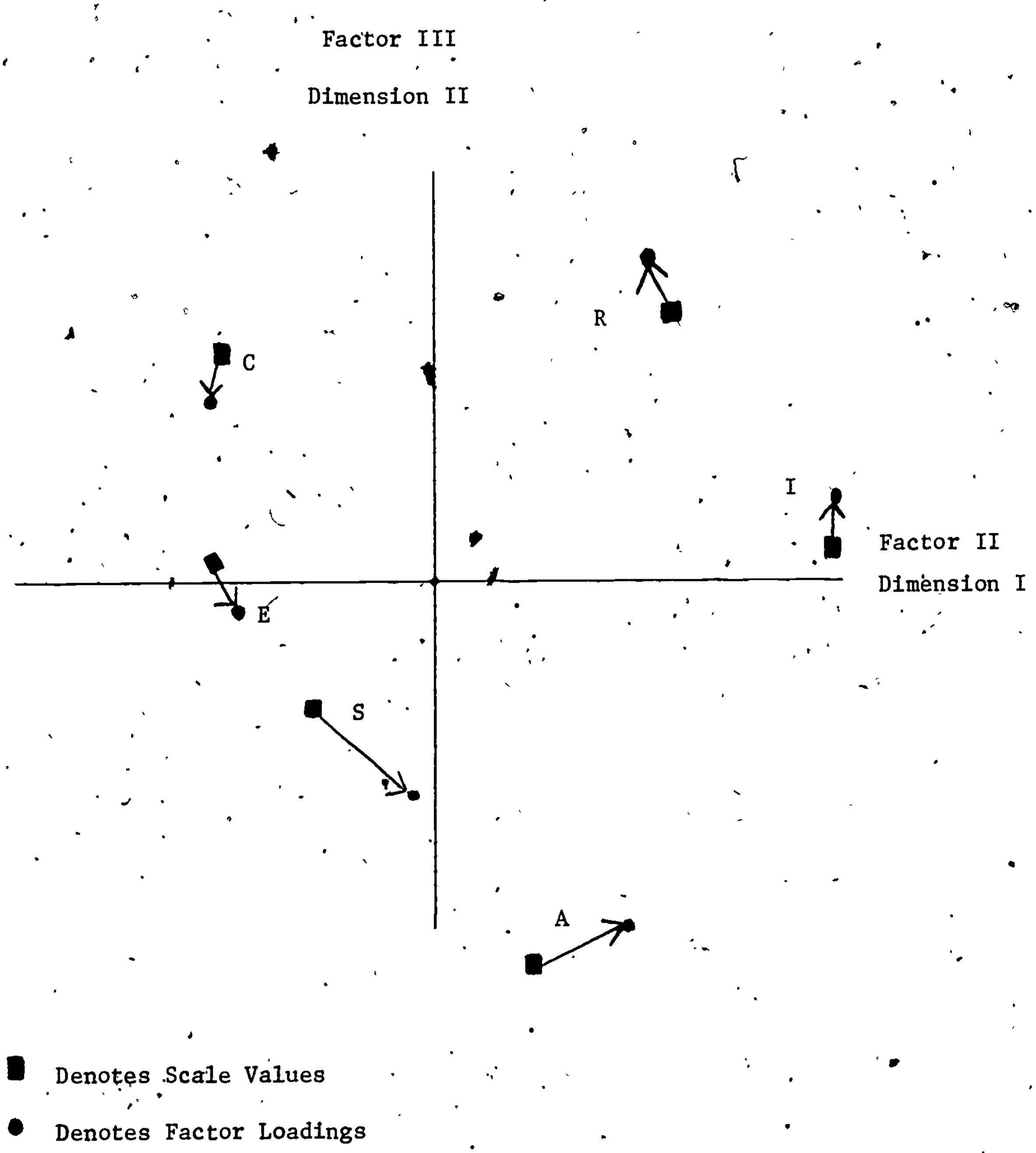

Figure 2. Scale Values and Factor Loadings for the <u>Vocational Preference Inventory</u> Scales

the corresponding scale value.

As in the first example above, one can compute a congruence coefficient for the two dimensions and the second and third principal components. That coefficient equals .97 and indicates that there is a high degree of correspondence between the 12 scale values along Dimensions I and II and the 12 loadings along Factors II and III.

Figure 2 graphically displays the relationship shown in Table 4 between Dimensions I and II of the scaling solution and Factors II and III of the components analysis. As in the earlier figure, squares represent subtest scale values, and circles represent subtest factor loadings. For each test, an arrow connects the square representing its scale values to the corresponding circle representing its factor loadings. This graph shows that the Factor II and III loadings (circles) place each interest subtest in the same region of the space as do its Dimension I and II scale values. Readers familiar with Holland's (1973) theory will recognize that the six squares (and the six circles) form the corners of a roughly hexagonal configuration, and the points fall along the hexagon in the order predicted by Holland's theory.

In Table 4 and Figure 2, we see the same kind of relationship between the multidimensional scaling and factor analysis that we saw in the earlier analysis of human ability data. That is, the first principal component has no counterpart in the scaling solution. Each component beyond the first, however, does have a counterpart among the multidimensional scaling dimensions.

Discussion

At this point, I must explain how the MDS scale values were obtained in the two examples. Two MDS dimensions were extracted using Kruskal, Young, and Seery's (1973) nonmetric MDS program KYST. The program rotated the two dimensions to what its authors call a principal components orientation. I then reflected dimensions as necessary so that the signs of scale values would more

closely approximate those of the corresponding factor loadings. Finally, I performed a uniform shrinking of the MDS dimensions. That is, the scale values were multiplied by a constant which was chosen so that the sum of squared scale values along Dimensions I and II would equal the sum of squared factor loadings along the second and third principal components.

This multiplication of the scale values by a constant is permissible because the origin and unit of measurement in a MDS are arbitrary. In factor analysis, on the other hand, the unit and origin for the loadings are fixed. Because they are fixed, the squared loading for a test can be interpreted as the proportion of variance in the test which is accounted for by the factor. No such proportion of variance interpretation can be made of the MDS scale values, because the unit and origin of those scale values are arbitrarily set.

Returning to the earlier examples, the GATB and VPI analyses, those analyses suggest the following two relationships between a MDS in K dimensions and a principal components analysis in K+1 factors. First, the components analysis will often contain a general factor with no counterpart among the scaling dimensions. Second, after applying an appropriate rotation and multiplicative constant to the MDS scale values, some or all of the components beyond the first will have a counterpart in the scaling solution.

The above two relationships deserve several caveats. In the interest of time, however, I will present only two. First, the above examples are only suggestive. I do not know how generally they hold. The generality of these two relationships deserves further investigation.

Second, the relationships can be expected to hold only when MDS and factor analysis are used to analyze the same intercorrelation matrix. Many of nonmetric MDS's advantages arise from the fact that it can be used to analyze data

for which conventional factor analysis is not considered appropriate. Researchers using MDS have often employed very different experimental procedures and have often analyzed very different data from those employed by factor analysts examining the same research question. The above relationships cannot be expected to describe the correspondence between a MDS and a factor analysis which are based on vastly different proximity data.

Within the limits set by these caveats, however, the above examples do suggest a correspondence between MDS and principal components analyses of the same test or test item intercorrelations. The factor solution is often more complex, in part because it contains a general factor with no counterpart among the scaling dimensions. In human ability data, the general component is sometimes called the general ability factor, and it has played a central role in several theories of human abilities. In self-report data, the general factor is sometimes called a response level component, and it has sometimes been totally ignored in reporting factor results. When the general factor is important, MDS omits a central feature of the test structure. When it is inconsequential, however, MDS may provide a simpler representation of the test structure which preserves all of its essential aspects.

References

Borman, W. C., Hough, L. M, & Dunnette, M. D. Development of behaviorally-based rating scales for evaluating the performance of U.S. Navy recruiters. (Technical Report NPRDC TR 76-31). San Diego, CA: Navy Personal Research and Development Center, 1976.

Cronbach, L. J. The two disciplines of scientific psychology. In D. N. Jackson & S. Messick (Eds.), Problems in human assessment. New York: McGraw-Hill, 1967, 22-39.

Hanson, G. R., Prediger, D. J., & Schussel, R. H. Development and validation of sex-balanced interest inventory scales. Iowa City, Iowa: The American College Testing Program, 1977 (ACT Research Report 78).

Holland, J. L. Manual for the Vocational Preference Inventory (6th ed.). Palo Alto, Calif.: Consulting Psychologists Press, 1965.

Holland, J. L. Making vocational choices: A theory of careers. Englewood Cliffs, NJ: Prentice-Hall, 1973.

Holland, J. L., Whitney, D. R., Cole, N. S., & Richards, J. M., Jr. An empirical occuaptional classification derived from a theory of personality and intended for practice and research (ACT Research Report No. 29). Iowa City, Iowa: The American College Testing Program, 1969.

Jones, L. E. & Young, F. W. Structure of a social environment: Longitudinal individual differences scaling of an intact group. Journal of Personality and Social Psychology, 1972, 24, 108-121.

Kruskal, J. B., Young, F. W., & Seery, J. B. How to use KYST, a very flexible program to do multidimensional scaling and unfolding. Unpublished manuscript, Bell Laboratories, 1973.

MacCallum, R. C. Relations between factor analysis and multidimensional scaling. Psychological Bulletin, 1974, 81, 505-516.

Richardson, M. W. Multidimensional psychophysics. Psychological Bulletin, 1938, 35, 659-660.

Rosenberg, S., & Jones, R. A method of investigating and representing a person's implicit theory of personality: Theodore Dreiser's view of people. Journal of Personality and Social Psychology, 1972, 22, 372-386.

Rounds, J. B., Jr., Davison, M. L., & Dawis, R. V. The fit between Strong-Campbell Interest Inventory general occupational themes and Holland's hexagonal model. Journal of Vocational Behavior, 1979, 15, 303-315.

Schlessinger, I. M. & Guttman, L. Smallest space analysis of intelligence and achievement tests. Psychological Bulletin, 1969, 71, 95-100.

Shepard, R. N. Introduction to volume I. In R. N. Shepard, A. K. Rommey, S. B. Nerlove (Eds.), Multidimensional scaling: Theory and applications in the behavioral sciences (Vol. 1). New York: Seminar Press, 1972.

Smith, R. J. & Siegal, A. I. A multidimensional scaling analysis of the job of civil defense director. Journal of Applied Psychology, 1967, 51, 476-480.

Spearman, C. "General intelligence" objectively determined and measured. American Journal of Psychology, 1904, 15, 201-293.

Tagiuri, R. T. Social preference and its perception. In R. Tagiuri & L. Petrullo (Eds.), Person perception and interpersonal behavior. Stanford, California: Stanford University Press, 1958.

Thurstone, L. L. Primary mental abilities. Chicago: University of Chicago Press, 1938.

U. S. Government Printing Office. Manual for the USTES. Washington D. C.: U. S. Department of Labor, Manpower Services, 1970.

Vernon, P. E. The structure of human abilities. New York: Wiley, 1950.

Young, G. & Householder, A. S. Discussion of a set of points in terms of their mutual distances. Psychometrika, 1938, 3, 19-22.